Calxeda

An introduction to Calxeda and Co-Design Opportunities for ARM-based Servers

SOS16

March 14, 2012

Karl Freund, VP Marketing
(karl.freund@calxeda.com)

Introducing Calxeda

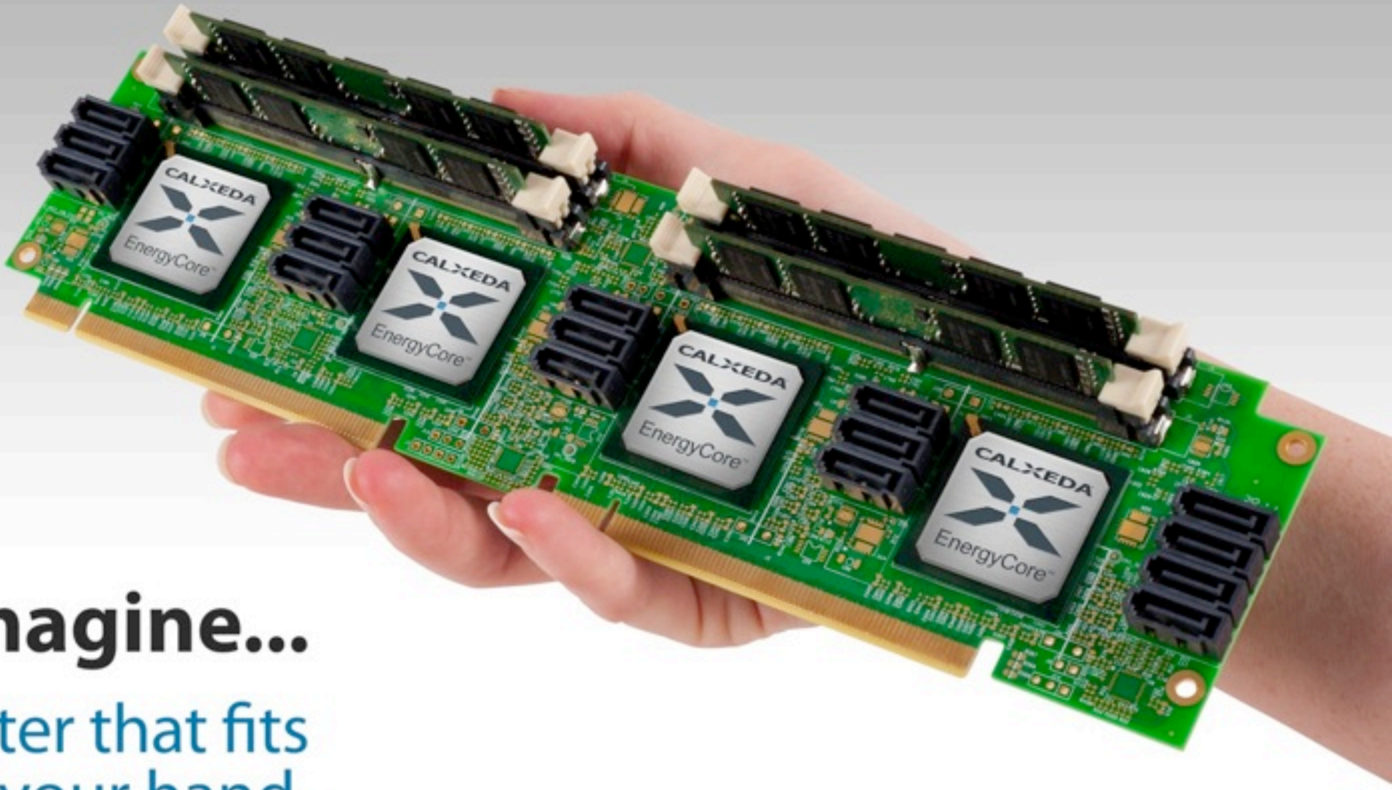
- **Creating the world's most efficient processor platform for hyper-scale server applications**

- **Austin-based and Venture-backed Semiconductor firm**
  - $48M in funding (August 2010)
    - Battery Ventures, Flybridge Capital, Highland Capital
    - ARM Holdings, ATIC strategic investments

- **Hewlett-Packard and other OEM's will ship servers this summer**

imagine...

A server that only uses
5 watts of energy!



imagine...

A server cluster that fits
in the palm of your hand.

(10x)

imagine...

Ten times the performance
at the same power
in the same space!

The Calxeda EnergyCore™ Processor SoC

Efficient

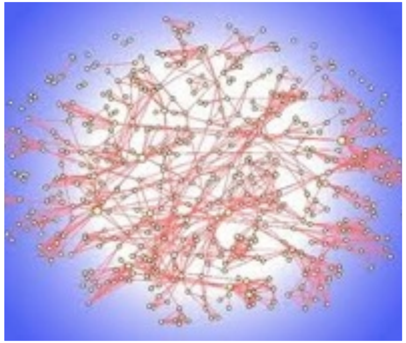
EnergyCore™

90% less energy

90% less space

50% lower costs

Scalable



Fabric Switch

Connects thousands of server nodes

Smart

Management

Autonomic power and system optimization

Targeted Application Characteristics

Applications built on portable or interpretive programming models (e.g., PHP, Ruby, Perl)

Scale out, parallel applications E.g., MapReduce, “Big Data”, Financial & Risk Modeling, “Data-Intensive HPC”)

Web Applications

Middle-Tier Applications

Offline Analytics

Storage/File Serving

Applications whose in-memory data domains are easily segmented into relatively small sets (e.g., memcached)

Applications that require high I/O throughput (e.g., Media streaming, Content Delivery, No-SQL databases)

Applications that are delivered via cloud infrastructures

EnergyCore architecture at a glance

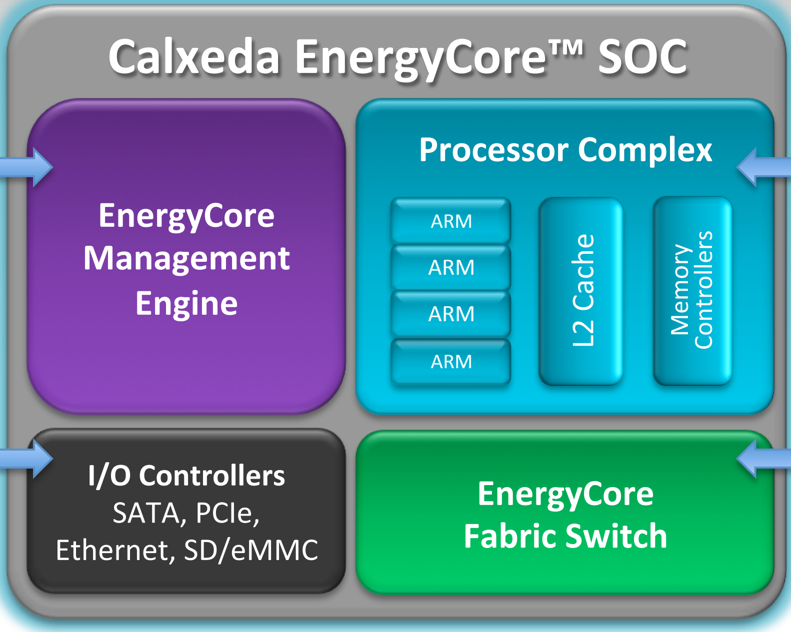
A complete building block for hyper-efficient computing

EnergyCore™ Management Engine

Advanced system, power and fabric optimization and management for energy-proportional computing

I/O Controllers

Standard drivers, standard interfaces. No surprises.



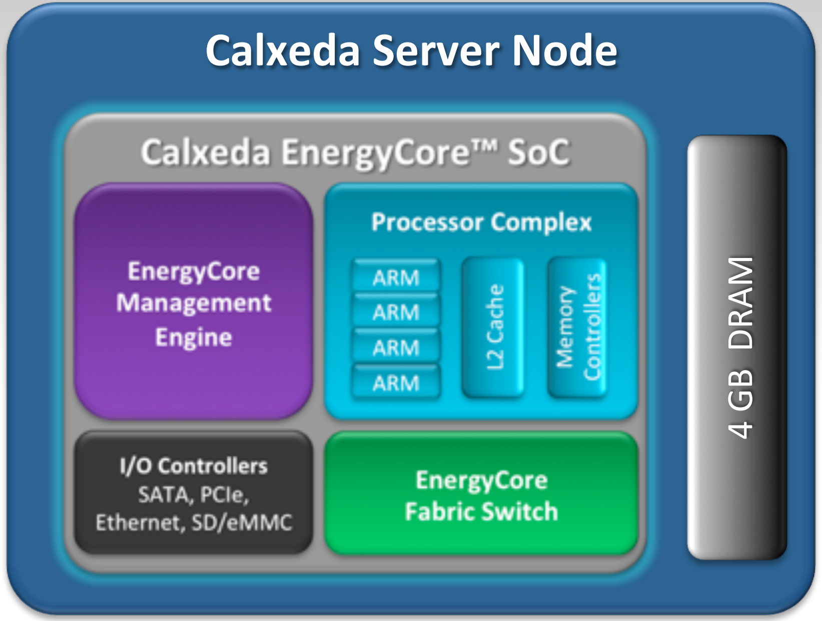
Processor Complex

Multi-core ARM® processors integrated with high bandwidth memory controllers

EnergyCore™ Fabric Switch

Integrated high-performance fabric converges internode communication, I/O, and storage networking

A Complete Server, only 5 Watts



Typical* Max Power: 5 Watts

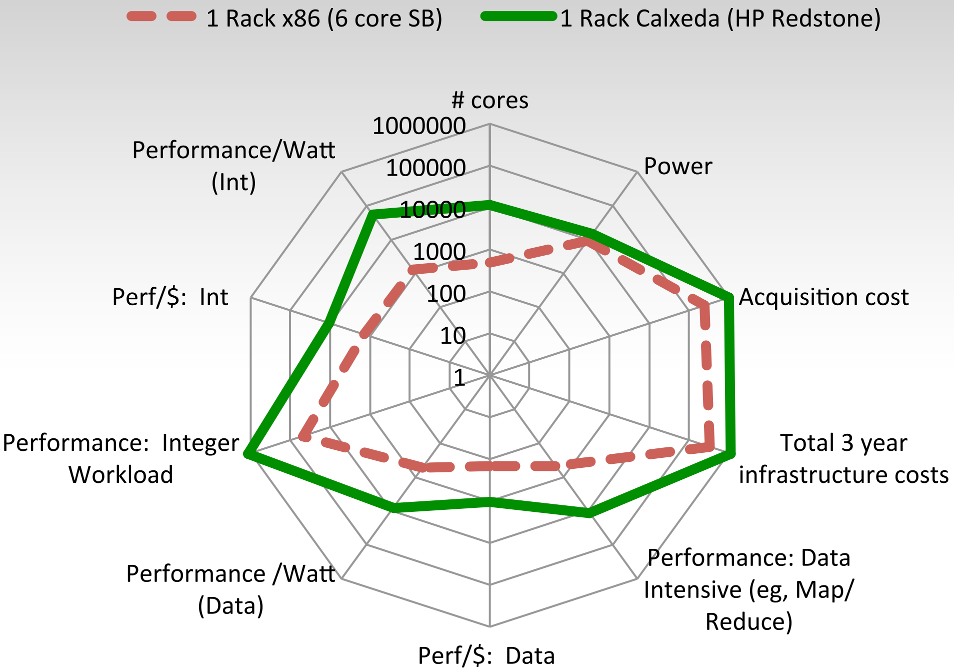
Power at Idle: < ½ Watt

* The power consumed under normal operating conditions under full application load (ie, 100% CPU utilization)

One Rack of Calxeda vs. One Rack of X86

2 Hypothetical Workloads:

1. Mix of I/O and Integer where Calxeda = 1/3 the per-core performance of Sandy Bridge
2. I/O Intensive such as Map/Reduce (1-1 ratio)



Breakthrough Savings and Simplicity

Energy, cost and space savings move the industry to new architecture

Traditional x86

$3.3M

400 servers
10 racks
20 switches
1,600 cables
91 kilowatts

89% less energy
94% less space
63% less cost
97% less complexity

HP 'Redstone' Development Platform

$1.2M

1,600 servers
1/2 rack
2 switches
41 cables
9.9 kilowatts

EnergyCard: a Quad-Node Reference Design

- Four-node reference platform from Calxeda
- Available as product and/or design
- Plugs into OEM system board with passive fabric, **no** additional switch HW
  EnergyCard delivers 80Gb Bandwidth to the system board. (8 x 10Gb links)

4 GB DRAM ECC mini-DIMMS
Quad-core servers
4 SATA / Node (flexibility!)
Power, SATA, & Fabric

Approximately 10”

4 Servers. Complete.
Only 20W.

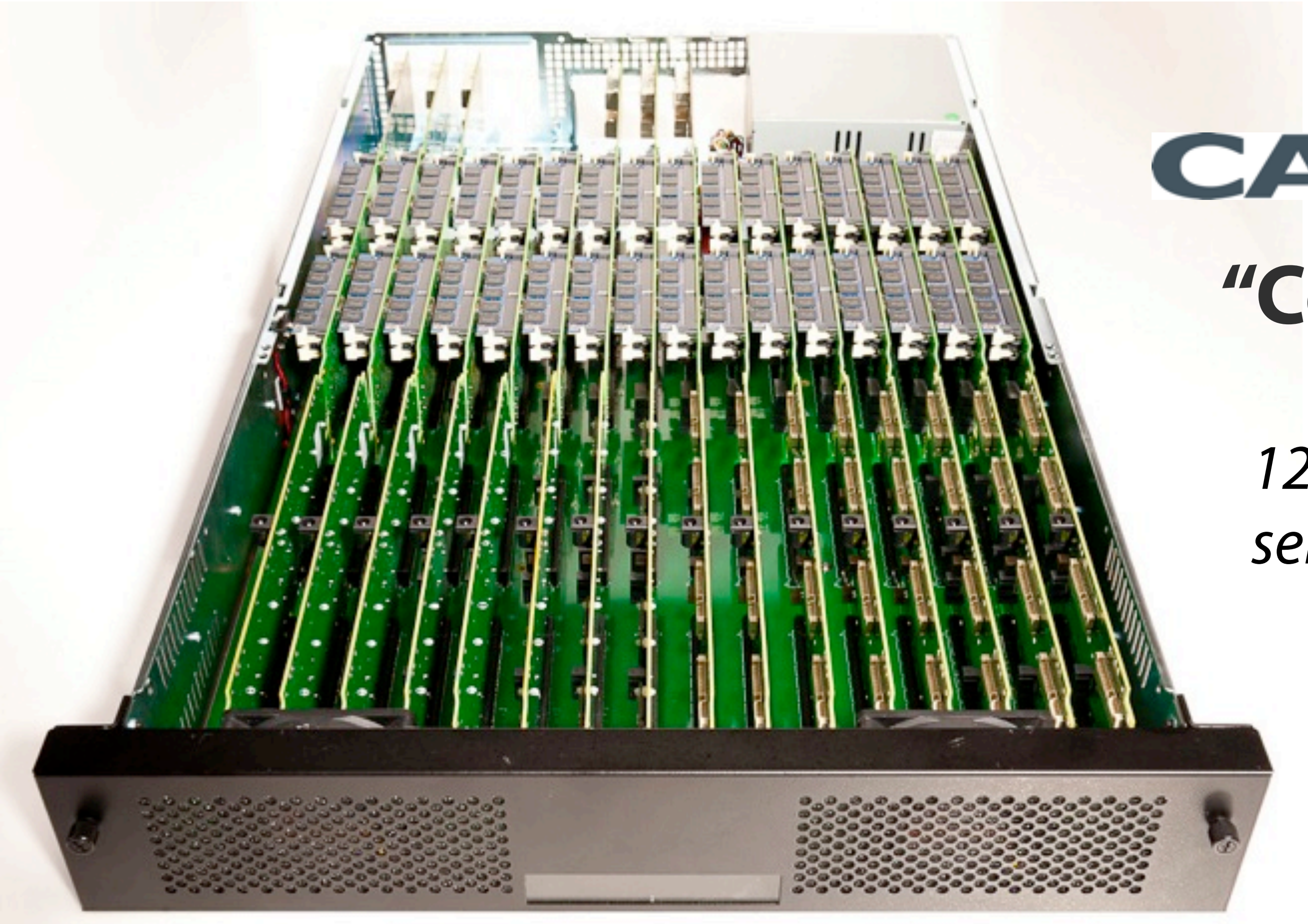
Redstone Development Platform

89% less energy

94% less space

64% less cost

288 EnergyCore server nodes in 4U



"Concept Car"

128 EnergyCore server nodes in 2U

Calxeda EnergyCards

Server Cluster Reference Designs

Calxeda Software Ecosystem – Base Packages (as of Feb 2012*)

Linux Kernel v3.2

ubuntu Server 12.04 LTS

Compilers/Languages
- GCC/gFortran 4.6.2
- PHP 5.3.8
- Perl 5.14.2
- Python 2.7.2, 3.2.2
- Ruby 1.8.7, 1.9.3
- Erlang r14

Debuggers/Profilers
- GDB 7.4
- GProf 2.13
- OProfile 0.9.6

Java
- Oracle JVM SEv7u4
- OpenJDK 6b24

Applications
- Apache 2.2.21
- Tomcat 6.0.32
- MySQL 5.5.17
- PostgreSQL 9.1

fedora v17+

Compilers/Languages
- GCC/gFortran 4.7.0
- PHP 5.4.0
- Perl 5.14.2
- Python 2.7.2, 3.2.2
- Ruby 1.8.7
- Erlang r14B

Debuggers/Profilers
- GDB 7.4
- GProf 2.13
- OProfile 0.9.6

Java
- Oracle JVM SEv7u4
- OpenJDK 6b24

Applications
- Apache 2.2.21
- Tomcat 7.0.25
- MySQL 5.5.20
- PostgreSQL 9.1.2

* Version numbers subject to change and are highly dependent on Linux distribution

Calxeda Software Ecosystem – HPC Packages (as of Feb 2012*)

Linux Kernel v3.2

ubuntu Server 12.04 LTS	fedora v17+
MPI • MPICH 1.2.7 • OpenMPI 1.4.3 • MPICH2 1.4.1 • Open-MX 3.5 **Checkpoint** • DMTCP 1.2.1 • Condor 7.2.4 **Libraries** • BLAS 1.2 • FFTW 2.1.5 • ScaLAPACK 1.8.0 **Monitoring** • Ganglia 3.1.7	**MPI** • ~~MPICH 1.2.7~~ • OpenMPI 1.5+ • MPICH2 1.4.1+ • ~~Open-MX 3.5~~ **Checkpoint** • ~~DMTCP 1.2.1~~ • Condor 7.4.2+ **Libraries** • ~~BLAS 1.2~~ • FFTW 3.3 • ScaLAPACK 1.7.5+ **Monitoring** • Ganglia 3.1.7

* Version numbers subject to change and are highly dependent on Linux distribution

Calxeda Software Ecosystem – Application Packages (as of Feb 2012*)

Linux Kernel v3.2	
ubuntu Server 12.04 LTS	fedora v17+
• **Apache Cassandra 1.0.7+** Packages to be provided by DataStax • **Apache Hadoop 1.0.0+** Packaged to be provided by Cloudera • **Memcached v1.4.13+**	• **Apache Cassandra 1.0.7+** Packages to be provided by DataStax • **Apache Hadoop 1.0.0+** Packaged to be provided by Cloudera • **Memcached v1.4.13+**

* Version numbers subject to change and are highly dependent on Linux distribution

Co-Design in the (ARM) SOC World

- **ARM needs help to be ready for HPC. And ARM has to care.**
  - Software infrastructure
  - Processor Core (e.g., ARM)
  - Other IP (DDR, Fabric, …)
- **SOC co-design represents opportunity**
  - Lower bar than full processor development investments
  - But still need to be broadly applicable
    - E.g.: Big Data analytics x HPC intersection
  - Examples:
    - Interconnect
    - Accelerators (e.g., GPUs)
    - Memory
    - I/O
    - Management

Calxeda Opportunities for Co-Design

1. High scale Interconnect designs
2. Management at Extreme Scale
3. Messaging optimizations
4. Heterogeneous computing
5. On-chip photonics
6. Resiliency
7. Memory technologies
8. Holistic power and performance optimizations

A note on performance and power estimates and comparisons

Calxeda comparisons are based on estimates of our power consumption and performance for **<u>appropriate</u>** workloads, i.e., appropriate workloads are typically memory-, I/O-, or network-bound when run on a single guest image per multi-core processor. Our power numbers are for "Typical Max", meaning 100% utilization in normal operating conditions. (The processor can also be capped at this or lower levels.)

For appropriate workloads, Calxeda performance will range from **1.3 – 3 Calxeda EnergyCore nodes (@4 core, 1.1 GHz) to one Intel Xeon E-5620 (4 core, 2.2Ghz)**. Hadoop, for example, is typically I/O bound and will be at the lower (better) end of the range. Static web hosting will be ~2.5-1. More integer-intensive applications are in the 3-1 range, while double precision floating point without tuning or SIMD exploitation will be higher (worse).

Assumptions in "Savings" calculations:

Performance ratios: (# of EnergyCore chips vs. E- 5620)

- Hadoop and Memcached: 1.5 - 1
- Static Web : 2.5 - 1
- Dynamic Web: 2.5 -1

PUE = 2.0

Power = $.10 / KWHr

Storage:

- Hadoop: 1 x 256GB SATA SSD/Node
- Static Web and memcached: Locally Diskless (ie, NAS or SAN)
- Dynamic Web: 1 x 1TB SATA / Node

Costs:

- 3 Year capital depreciation of networking and servers
- Direct and indirect power costs
- Does not include Datacenter Capital Depreciation space savings